FSF3847
Convex Optimization
with Engineering Applications
Assignment 3

Franco Ruggeri

May 12, 2023

1 Exercise 3.1

Problem

A simple example. Consider the optimization problem

$$\begin{array}{ll} \text{minimize} & x^2 + 1 \\ \text{subject to} & (x-2)(x-4) \leq 0, \end{array} \tag{P}$$

with variable $x \in \mathbb{R}$.

(a) *Analysis of primal problem.* Give the feasible set, the optimal value, and the optimal solution.

(b) *Lagrangian and dual function.* Plot the objective $x^2 + 1$ versus x. On the same plot, show the feasible set, optimal point and value, and plot the Lagrangian $L(x, \lambda)$ versus x for a few positive values of λ. Verify the lower bound property ($p^* \geq \inf_x L(x, \lambda)$ for $\lambda \geq 0$. Derive and sketch the Lagrange dual function g.

(c) *Lagrangian dual problem.* State the dual problem, and verify that it is a concave maximization problem. Find the dual optimal value and dual optimal solution λ^*. Does strong duality hold?

(d) *Sensitivity analysis.* Let $p^*(u)$ denote the optimal value of the problem

$$\begin{array}{ll} \text{minimize} & x^2 + 1 \\ \text{subject to} & (x-2)(x-4) \leq u, \end{array} \tag{P$_\mathrm{u}$}$$

as a function of the parameter u. Plot $p^*(u)$. Verify that $dp^*(0)/du = -\lambda^*$.

Solution

Part (a)

The feasible set $\mathcal{F}$ is:

$$\begin{aligned}\mathcal{F} &= \{x \in \mathbb{R} \mid (x-2)(x-4) \leq 0\} \\ &= \{x \in \mathbb{R} \mid 2 \leq x \leq 4\}\end{aligned}$$

The objective function $f_o(x) \coloneqq x^2+1$ is differentiable and has positive derivative in F:

$$\frac{\mathrm{d}f_o(x)}{\mathrm{d}x} = 2x > 0 \iff x \in \mathbb{R}_{++} \supset \mathcal{F}$$

which means $f_o(x)$ is monotonically increasing in $\mathcal{F}$. Thus, the optimal solution x^* and the optimal value p^* are:

$$x^* = \min \mathcal{F} = 2 \tag{1}$$

$$p^* = f_o(x^*) = 5 \tag{2}$$

These results can also be seen graphically in Fig. 1.

Part (b)

Let $f_1(x) \coloneqq (x-2)(x-4)$ be the constraint in the primal problem (P). The Lagrangian $L(x, \lambda)$ is:

$$\begin{aligned}L(x,\lambda) &= f_o(x) + \lambda f_1(x) \\ &= x^2 + 1 + \lambda(x-2)(x-4) \\ &= (1+\lambda)x^2 - 6\lambda x + (1+8\lambda) \\ &= (1+\lambda)\left(x - \frac{3\lambda}{1+\lambda}\right)^2 + 1 + 8\lambda - \frac{9\lambda^2}{1+\lambda}\end{aligned}$$

Hence, $L(x,\lambda)$ is a translated parabola and finding the infimum is easy. When $1+\lambda > 0$, $L(x,\lambda)$ is strictly convex and the minimum is $x^* = \frac{3\lambda}{1+\lambda}$. When $1+\lambda < 0$, $L(x,\lambda)$ is concave and unbounded below. For $\lambda = -1$, $L(x,\lambda)$ is undefined, but we write its infimum to be $-\infty$ to simplify notation. Thus, the Lagrangian dual function $g(\lambda)$ is:

$$g(\lambda) = \inf_x L(x,\lambda) = \begin{cases} 1 + 8\lambda - \frac{9\lambda^2}{1+\lambda} & \text{if } \lambda > -1 \\ -\infty & \text{otherwise} \end{cases} \tag{3}$$

Fig. 2 illustrates the function for $\lambda \geq 0$, which are the feasible points for the dual problem. Using Eqs. (2) and (3), we can verify the lower bound property

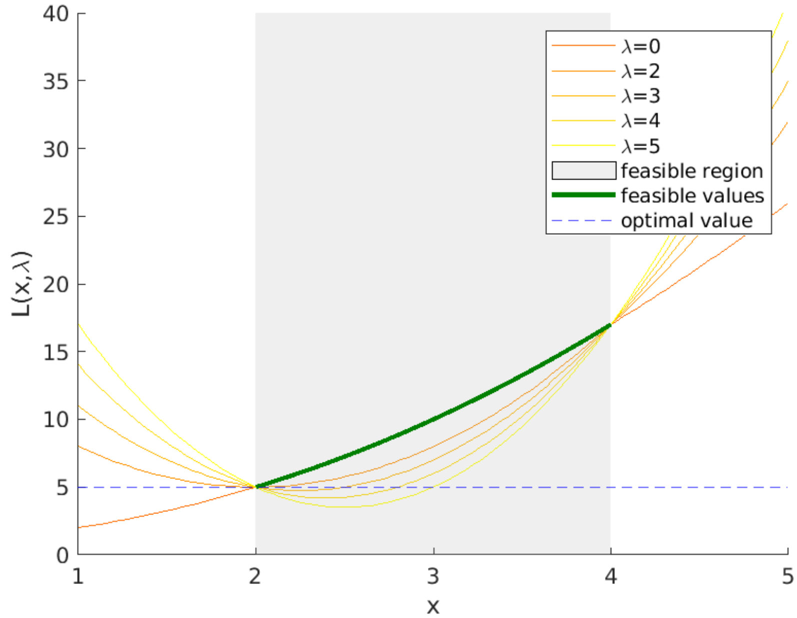

Figure 1: Lagrangian $L(x, \lambda)$ versus x for a few positive values of λ. In the feasible set, the Lagrangian $L(x, \lambda)$ is always an underestimate of the objective function $f_o(x) = L(x, 0)$.

$p^* \geq \inf_x L(x, \lambda)$:

$$\begin{aligned}
g(\lambda) = \inf_x L(x, \lambda) \leq p^* &\iff \\
1 + 8\lambda - \frac{9\lambda^2}{1+\lambda} \leq 5 &\iff \\
\frac{(\lambda-2)^2}{\lambda+1} \geq 0 &\iff \\
\lambda \in (-1, \infty) \supset [0, \infty] &
\end{aligned}$$

This result can also be observed graphically in Fig. 1, which shows that the Lagrangian $L(x, \lambda)$ is always an underestimate of the objective function $f_o(x)$ in the feasible set. Such a property was expected, since minimizing the Lagrangian is a relaxation of the primal problem (P) [1].

Part (c)

The dual problem of (P) is:

$$\begin{aligned}
&\text{maximize} && g(\lambda) \\
&\text{subject to} && \lambda \geq 0
\end{aligned} \tag{D}$$

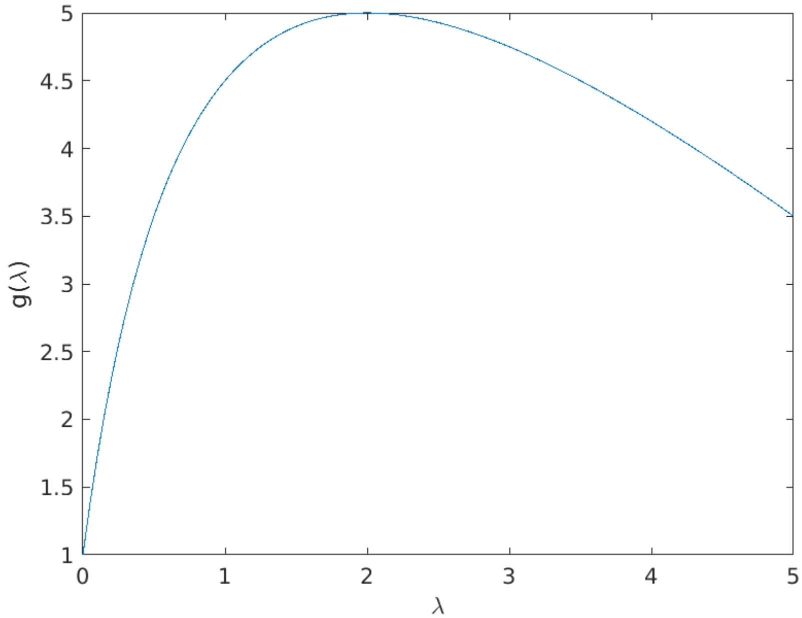

Figure 2: Lagrangian dual function $g(\lambda) = \inf_x L(x, \lambda)$ for $\lambda \geq 0$. It can be immediately seen that $g(\lambda)$ is concave for $\lambda \geq 0$, the maximum value is $\lambda^* = 2$.

The Lagrangian dual function $g(\lambda)$ is concave in λ, since it is the infimum of the Lagrangian $L(x, \lambda)$, which is affine (hence, concave) in λ, and the infimum preserves concavity [1]. This means that local maxima are also global maxima. The Lagrangian dual function $g(\lambda)$ is differentiable in the feasible set ($\lambda \geq 0$) and has a stationary point (minimum) in the feasible set, as can be found by setting the derivative to 0:

$$\frac{\mathrm{d}g(\lambda)}{\mathrm{d}\lambda} = 8 + \frac{-18\lambda(1+\lambda) + 9\lambda^2}{(1+\lambda)^2} = \frac{(\lambda+4)(\lambda-2)}{(1+\lambda)^2} = 0 \iff \lambda = 2$$

where the infeasible points ($\lambda < 0$) have been ignored. Thus, the optimal dual point and the optimal dual value are:

$$\begin{aligned} \lambda^* &= 2 \\ g(\lambda^*) &= 5 = p^* \end{aligned} \tag{4}$$

where Eq. (4) indicates that there is strong duality.

Part (d)

The feasible set $\mathcal{F}_u$ of the perturbated problem $(\mathrm{P_u})$ is:

$$\begin{aligned}\mathcal{F}_u &= \{x \in \mathbb{R} \mid (x-2)(x-4) \leq u\} \\ &= \{x \in \mathbb{R} \mid 3-\sqrt{1+u} \leq x \leq 3+\sqrt{1+u}\}\end{aligned}$$

where the square root $\sqrt{1+u}$ implies $u \geq -1$. For $u < -1$, $\mathcal{F}_u = \emptyset$ and the problem $(\mathrm{P_u})$ is infeasible (indicated by optimal value ∞). The upper bound of the feasible set is always positive (i.e., $3+\sqrt{1+u} > 0$). Thus, the optimal point $x^*(u)$ is:

$$\begin{aligned}x^*(u) &= \begin{cases} 0 & \text{if } 3-\sqrt{1+u} < 0 \\ 3-\sqrt{1+u} & \text{if } 3-\sqrt{1+u} > 0 \end{cases} \\ &= \begin{cases} 0 & \text{if } u \geq 8 \\ 3-\sqrt{1+u} & \text{if } -1 \leq u < 8 \end{cases}\end{aligned}$$

where $x = 0$ (minimum of $f_o(x)$) does not belong to the feasible set F_u in the first case, and does belong to the feasible set in the second case. The optimal value $p^*(u)$ is:

$$p^*(u) = \begin{cases} 1 & \text{if } u \geq 8 \\ 11 - 6\sqrt{1+u} + u & \text{if } -1 \leq u < 8 \\ \infty & \text{if } u < -1 \end{cases}$$

The function $p^*(u)$ is differentiable for $u \geq -1$:

$$\frac{\mathrm{d}p^*(u)}{\mathrm{d}u} = \begin{cases} 0 & \text{if } u \geq 8 \\ 1 - \frac{3}{\sqrt{1+u}} & \text{if } -1 \leq u < 8 \end{cases}$$

And we can verify that $\frac{\mathrm{d}p^*(0)}{\mathrm{d}u} = -2 = \lambda^*$. That is, the optimal dual solution describes the sensitivity of the optimal value to perturbations in the constraint.

2 Exercise 3.2

Problem

Robust linear programming with polyhedral uncertainty. Consider the robust LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \sup_{a \in \mathcal{P}_i} a^T x \leq b_i, \; i = 1, ..., m, \end{array} \tag{LP1}$$

with variable $x \in \mathbb{R}^n$, where $\mathcal{P}_i = \{a \mid C_i a \preceq d_i\}$. The problem data are $c \in \mathbb{R}^n$, $C_i \in \mathbb{R}^{m_i \times n}$, $d_i \in \mathbb{R}^{m_i}$, and $b \in \mathbb{R}^m$. We assume the polyhedra $\mathcal{P}_i$ are nonempty.

Show that this problem is equivalent to the LP

$$\begin{array}{ll}\text{minimize} & c^T x \\ \text{subject to} & d_i^T z_i \le b_i,\ i = 1, ..., m, \\ & C_i^T z_i = x,\ i = 1, ..., m, \\ & z_i \succeq 0,\ i = 1, ..., m\end{array} \tag{LP2}$$

with variables $x \in \mathbb{R}^n$ and $z_i \in \mathbb{R}^{m_i}$, $i = 1, ..., m$. *Hint.* Find the dual of the problem of maximizing $a_i^T x$ over $a_i \in \mathcal{P}_i$ (with variable a_i).

Solution

Let $f_i(x) \coloneqq \sup_{a \in \mathcal{P}_i} a^T x$ be the left-hand side of the constraint in (LP1). The function $f_i(x)$ can be written as a LP:

$$\begin{array}{ll}\underset{a \in \mathbb{R}^n}{\text{maximize}} & a^T x \\ \text{subject to} & C_i a - d_i \preceq 0,\end{array}$$

i.e., $f_i(x) = \text{optval}(CLP)$. Introducing a slack variable, this problem can be transformed to a standard dual linear program:

$$\begin{array}{ll}\underset{a \in \mathbb{R}^n}{\text{maximize}} & x^T a \\ \text{subject to} & C_i a + s = d_i, \\ & s \succeq 0\end{array}$$

Using the duality result, the corresponding primal linear program is:

$$\begin{array}{ll}\underset{z_i \in \mathbb{R}^n}{\text{minimize}} & d_i^T z_i \\ \text{subject to} & C_i^T z_i = x, \\ & z_i \succeq 0\end{array} \tag{CPLP}$$

Let $\mathcal{F}_i \coloneqq \{z_i \mid C_i^T z_i = x, z_i \succeq 0\}$ be the feasibility set of (CPLP). Since linear programs have strong duality, we have:

$$\text{optval}(CLP) = \text{optval}(CDLP) \iff \sup_{a \in \mathcal{P}_i} a^T x = \inf_{z_i \in \mathcal{F}_i} d_i^T z_i$$

Thus, the original problem (LP1) is equivalent to:

$$\begin{array}{ll}\underset{x \in \mathbb{R}^n}{\text{minimize}} & c^T x \\ \text{subject to} & \inf_{z_i \in \mathcal{F}_i} d_i^T z_i \le b_i,\ i = 1, ..., m,\end{array} \tag{LP3}$$

Observe that b_i is greater than the infimum of a set of elements if and only there exists one element of the set which is smaller than b_i:

$$b_i \geq \inf_{z_i \in \mathcal{F}_i} d_i^T z_i \iff \exists z_i \in \mathcal{F}_i : b_i \geq d_i^T z_i \tag{5}$$

Therefore, the problem is equivalent to:

$$\begin{array}{ll} \underset{x,\, z_i \,\in\, \mathbb{R}^n}{\text{minimize}} & c^T x \\ \text{subject to} & d_i^T z_i \leq b_i,\ i = 1, ..., m, \\ & C_i^T z_i = x,\ i = 1, ..., m, \\ & z_i \succeq 0,\ i = 1, ..., m \end{array}$$

which is exactly (LP2). The equivalence can be understood as follows. Assume that, in (LP2), we optimize first over z_i and then over x, for $i = 1, ..., m$. While optimizing over z_i, x is considered fixed. Hence, the objective function is constant and the optimization over z_i is a feasibility problem of finding at least one $z_i \in \mathcal{F}_i$ satisfying the constraint $b_i \geq d_i^T z_i$. From Eq. (5), this feasibility problem is equivalent to satisfying the constraint in (LP3).

3 Exercise 3.3

Problem

Dual of channel capacity problem. Derive a dual for the problem

$$\begin{array}{ll} \text{minimize} & -c^T x + \sum_{i=1}^{m} y_i \log y_i \\ \text{subject to} & Px = y, \\ & x \succeq 0, \\ & \mathbf{1}^T x = 1 \end{array} \tag{P}$$

where $P \in \mathbb{R}^{m \times n}$ has nonnegative elements, and its columns add up to one (i.e., $P^T \mathbf{1} = \mathbf{1}$). The variables are $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$. Simplify the dual problem as much as possible.

Solution

Starting from (P), rewrite the constraints as follows:

$$\begin{array}{ll} \underset{x \,\in\, \mathbb{R}^n,\, y \,\in\, \mathbb{R}^m}{\text{minimize}} & -c^T x + \sum_{i=1}^{m} y_i \log y_i \\ \text{subject to} & Px - y = 0, \\ & x \succeq 0, \\ & \mathbf{1}^T x - 1 = 0 \end{array}$$

Let $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}^n_+$, $\nu \in \mathbb{R}$ be the Lagrangian multipliers. The Lagrangian $L(x, y, \lambda, \mu, \nu)$ is:

$$\begin{aligned}
L(x, y, \lambda, \mu, \nu) &= -c^T x + \sum_{i=1}^{m} y_i \log y_i + \lambda^T (Px - y) - \mu^T z + \nu(\mathbf{1}^T z + 1) \\
&= \underbrace{\left(-c + P^T \lambda - \mu + \nu \mathbf{1}\right)^T z}_{f_1(x,\lambda,\mu,\nu):=} + \underbrace{\sum_{i=1}^{m} y_i (\log y_i - \lambda_i)}_{f_2(y,\lambda):=} + \nu
\end{aligned}$$

The function f_1 is linear in x, hence unbounded below unless $-c+P^T\lambda-\mu+\nu\mathbf{1} = 0$. The function f_2 is convex and differentiable, so the minimum point y^* and the minimum value $f_2(y^*)$ can be found by setting the partial derivatives to 0:

$$\left.\frac{\partial f_2(y, \lambda)}{\partial y_i}\right|_{y=y^*} = \log y_i^* + 1 - \lambda_i = 0 \iff y_i^* = e^{\lambda_i - 1}$$

$$f_2(y^*, \lambda) = \sum_{i=1}^{m} e^{\lambda_i - 1} (\log e^{\lambda_i - 1} - \lambda_i) = -\sum_{i=1}^{m} e^{\lambda_i - 1}$$

Therefore, the Lagrangian dual function $g(\lambda, \mu, \nu)$ is:

$$g(\lambda, \mu, \nu) = \begin{cases} -\sum_{i=1}^{m} e^{\lambda_i - 1} - \nu & \text{if } -c + P^T \lambda - \mu + \nu \mathbf{1} = 0 \\ -\infty & \text{otherwise} \end{cases}$$

and the dual problem is:

$$\begin{aligned}
\underset{\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^n, \nu \in \mathbb{R}}{\text{maximize}} \quad & -\sum_{i=1}^{m} e^{\lambda_i - 1} - \nu \\
\text{subject to} \quad & -c + P^T \lambda - \mu + \nu \mathbf{1} = 0, \\
& \mu \succeq 0
\end{aligned}$$

Eliminate the slack variable μ:

$$\begin{aligned}
\underset{\lambda \in \mathbb{R}^m, \nu \in \mathbb{R}}{\text{maximize}} \quad & -\sum_{i=1}^{m} e^{\lambda_i - 1} - \nu \\
\text{subject to} \quad & -c + P^T \lambda + \nu \mathbf{1} \succeq 0,
\end{aligned}$$

The constraint can be simplified by using the property $P^T \mathbf{1} = \mathbf{1}$:

$$\begin{aligned}
-c + P^T(\lambda + \nu\mathbf{1} - \nu\mathbf{1}) + \nu\mathbf{1} \succeq 0 &\iff \\
-c + P^T(\lambda + \nu\mathbf{1}) - \nu P^T\mathbf{1} + \nu\mathbf{1} \succeq 0 &\iff \\
P^T(\lambda + \nu\mathbf{1}) - c \succeq 0 &
\end{aligned}$$

Thus, the dual problem is equivalent to:

$$\begin{array}{ll} \underset{\lambda \in \mathbb{R}^m, \nu \in \mathbb{R}}{\text{maximize}} & -\sum_{i=1}^{m} e^{\lambda_i - 1} - \nu \\ \text{subject to} & P^T(\lambda + \nu \mathbf{1}) - c \succeq 0, \end{array}$$

The constraint can be written as a standard inequality ($Ax - b \succeq 0$) with a variable change $\omega \coloneqq \lambda + \nu\mathbf{1}$:

$$\begin{array}{ll} \underset{\omega \in \mathbb{R}^m, \nu \in \mathbb{R}}{\text{maximize}} & - e^{-\nu-1} \sum_{i=1}^{m} e^{\omega_i} - \nu \\ \text{subject to} & P^T\omega - c \succeq 0, \end{array} \tag{P2}$$

Now, let $f_o(\omega, \nu) \coloneqq -e^{-\nu-1} \sum_{i=1}^{m} e^{\omega_i} - \nu$ be the objective function. Since (P2) has no constraints on ν, we can easily solve the *unconstrained partial* optimization over ν by setting the partial derivative to 0:

$$\left.\frac{\partial f_o(\omega, \nu)}{\partial \nu}\right|_{\nu=\nu^*} = e^{-\nu^*-1} \sum_{i=1}^{m} e^{\omega_i} - 1 = 0 \iff \nu^* = \log \sum_{i=1}^{m} e^{\omega_i - 1}$$

$$f_o(\omega, \nu^*) = -1 - \log \sum_{i=1}^{m} e^{\omega_i - 1}$$

Therefore, the simplified dual problem is:

$$\begin{array}{ll} \underset{\omega \in \mathbb{R}^m}{\text{maximize}} & -1 - \log \sum_{i=1}^{m} e^{\omega_i - 1} \\ \text{subject to} & P^T\omega - c \succeq 0, \end{array}$$

4 Exercise 3.4

Problem

SDP relaxations of two-way partitioning problem. We consider the two-way partitioning problem

$$\begin{array}{ll} \text{minimize} & x^T W x \\ \text{subject to} & x_i^2 = 1, \; i = 1, ..., n, \end{array} \tag{P1}$$

with variable $x \in \mathbb{R}^n$. The Lagrange dual of this (nonconvex) problem is given by the SDP

$$\begin{array}{ll} \text{maximize} & -\mathbf{1}^T\nu \\ \text{subject to} & W + \mathbf{diag}(\nu) \succeq 0, \end{array} \tag{SDP1}$$

with variable $\nu \in \mathbb{R}^n$. The optimal value of this SDP gives a lower bound on the optimal value of the partitioning problem. In this exercise, we derive another SDP that gives a lower bound on the optimal value of the two-way partitioning problem, and explore the connection between the two SDPs.

(a) *Two-way partitioning problem in matrix form.* Show that the two-way partitioning problem can be cast as

$$\begin{array}{ll} \text{minimize} & \mathbf{trace}(WX) \\ \text{subject to} & X \succeq 0, \\ & \mathbf{rank}(X) = 1, \\ & X_{ii} = 1,\ i = 1, ..., n \end{array} \tag{P2}$$

with variable $X \in \mathbb{S}^n$. *Hint.* Show that if X is feasible, then it has the form $X = xx^T$, where $x \in \mathbb{R}^n$ satisfies $x_i \in \{-1, 1\}$ (and vice versa).

(b) *SDP relaxation of two-way partitioning problem.* Using the formulation in part (a), we can form the relaxation

$$\begin{array}{ll} \text{minimize} & \mathbf{trace}(WX) \\ \text{subject to} & X \succeq 0, \\ & X_{ii} = 1,\ i = 1, ..., n \end{array} \tag{SDP2}$$

with variable $X \in \mathbb{S}^n$. This problem is an SDP, and therefore can be solved efficiently. Explain why its optimal value gives a lower bound on the optimal value of the two-way partitioning problem (P1). What can you say if an optimal point X^* for this SDP has rank one?

(c) We now have two SDPs that give a lower bound on the optimal value of the two-way partitioning problem (P1): the SDP relaxation (SDP2) found in part (b), and the Lagrange dual of the two-way partitioning problem, given in (SDP1). What is the relation between the two SDPs? What can you say about the lower bounds found by them? *Hint*: Relate the two SDPs via duality.

Remark: Note that if $M \in \mathbb{R}^{n\times n}$, $M = M^T$, and $x \in \mathbb{R}^n$, then $x^TMx = \mathbf{trace}(x^TMx) = \mathbf{trace}(Mxx^T)$.

Solution

Part (a)

The objective function of (P1) can be re-written as:

$$x^T W x = \textbf{trace}(x^T W x) = \textbf{trace}(W x x^T)$$

Introducing $X = xx^T$, the problem (P1) is equivalent to:

$$\begin{aligned} &\text{minimize} && \textbf{trace}(WX) \\ &\text{subject to} && X = xx^T, \\ & && x_i^2 = 1,\ i = 1, ..., n \end{aligned}$$

Thus, (P1) is equivalent to (P2) if and only if their constraints are equivalent. That is, we have to prove that:

$$\begin{cases} X = xx^T \\ x_i^2 = 1,\ i = 1, ..., n \end{cases} \iff \begin{cases} X \succeq 0 \\ \textbf{rank}(X) = 1 \\ X_{ii} = 1,\ i = 1, ..., n \end{cases}$$

First, we prove the "right implication" $\implies$. Given $X = xx^T$ and $x_i^2 = 1$ for $i = 1, ..., n$, we have:

- $X_{ii} = 1$: By construction, $X_{ij} = x_i x_j$, hence $X_{ii} = x_i^2 = 1$.
- X is symmetric: $X^T = (xx^T)^T = xx^T = X$.
- $X \succeq 0$: $X_{ii} = 1$ and $X_{ij} \pm 1$, so X is diagonally dominant, which implies X is positive semidefinite.
- $\textbf{rank}(X) = 1$: $\textbf{rank}(X) = \textbf{rank}(xx^T) = \textbf{rank}(x) = 1$, as $x \in \mathbb{R}^n$.

Second, we prove the "left implication" $\impliedby$. Given $X \succeq 0$, $\textbf{rank}(X) = 1$, $X_{ii} = 1$ for $i = 1, ..., n$, we have:

- $X = xx^T$: Since $\textbf{rank}(X) = \textbf{dim}(\textbf{Col}(X) = 1$, the columns of X are multiple of some vector $v \in \mathbb{R}^n$, i.e., $X = \begin{bmatrix} u_1 v & ... & u_n v \end{bmatrix}$. Thus, $X = uv^T$ is the product of two vectors[1]. Let $\hat{u} := \frac{u}{\|u\|}$ and $\hat{v} := \frac{v}{\|v\|}$ be the normalized

[1] Inspired by https://math.stackexchange.com/questions/1545118/a-rank-one-matrix-is-the-product-of-two-vectors.

vectors u and v, respectively. Since X is symmetric, we have[2]:

$$\begin{aligned}
X &= X^T && \iff \\
uv^T &= (uv^T)^T && \iff \\
\|u\|\|v\|\hat{u}\hat{v}^T &= \|u\|\|v\|(\hat{u}\hat{v}^T)^T && \iff \\
\hat{u}\hat{v}^T &= (\hat{u}\hat{v}^T)^T && \iff \\
\hat{u}\hat{v}^T &= \hat{v}\hat{u}^T && \iff \\
\hat{u}^T\hat{u}\hat{v}^T\hat{v} &= \hat{u}^T\hat{v}\hat{u}^T\hat{v} && \iff \\
1 &= (\hat{u}^T\hat{v})^2 && \iff \\
1 &= |\hat{u}^T\hat{v}| && \iff \\
\|\hat{u}\|\|\hat{v}\| &= |\hat{u}^T\hat{v}|
\end{aligned}$$

where the last equality is the equality case of Cauchy–Schwarz inequality and holds only if $\hat{u}$ and $\hat{v}$ are linearly dependent. Thus, $u = \lambda v$ for some $\lambda \in \mathbb{R}$ and $X = \lambda u u^T$. Let $x \coloneqq \sqrt{\lambda} u$. We have found that $X = xx^T$.

- $x_i^2 = 1$ for $i = 1, ..., n$: By construction, $X_{ij} = x_i x_j$, hence $x_i^2 = X_{ii} = 1$.

Part (b)

Let $f_o(X) \coloneqq \mathbf{trace}(WX)$ be the the objective function in (P2) and (SDP2). Let $\mathcal{F}$ and $\mathcal{F}_R$ be the feasible set of (P2) and (SDP2), respectively:

$$\begin{aligned}
\mathcal{F}_R &\coloneqq \{X \mid X \succeq 0, X_{ii} = 1,\ i = 1, ..., n\} \\
\mathcal{F} &\coloneqq \mathcal{F}_R \cap \{X \mid \mathbf{rank}(X) = 1\} \subseteq \mathcal{F}_R
\end{aligned}$$

Let $X^* \in \mathcal{F}$ and $X_R^* \in \mathcal{F}_R$ be the optimal solutions to (P2) and (SDP2), respectively. By definition of optimal value, we have:

$$\begin{aligned}
& f_o(X_R^*) \le f_o(X),\ \forall X \in \mathcal{F}_R \\
\implies & f_o(X_R^*) \le f_o(X),\ \forall X \in \mathcal{F} \\
\implies & f_o(X_R^*) \le f_o(X^*)
\end{aligned} \tag{6}$$

Now, assume $\mathbf{rank}(X_R^*) = 1$, which means the optimal solution lies in the feasible set of (P2) (i.e., $X_R^* \in \mathcal{F}$). Then, Eq. (6) becomes exactly the definition of optimal value of (P2). Thus, X_R^* and $f_o(X_R^*)$ are optimal solution and optimal value for (P2), respectively. This happens because the objective function $f_o(X)$ is the same for (P) and (SDP2). Intuitively, if the optimal solution to the relaxation is *not* feasible for the original problem ($X_R^* \in \mathcal{F}_\mathcal{R} \setminus \mathcal{F}$), the optimal value of the relaxation gives a lower bound. Instead, if the optimal solution to the relaxation is feasible for the original problem ($X_R^* \in \mathcal{F}$), then it is also an optimal solution for the original problem.

[2] Inspired by https://math.stackexchange.com/questions/359604/form-of-symmetric-matrix-of-rank-one.

Part (c)

Starting from problem (SDP1), let $\Lambda \in \mathbb{S}_+^n$ be the Lagrangian multiplier. The Lagrangian $L(\nu, \Lambda)$ is:

$$\begin{aligned} L(\nu, \Lambda) &= \mathbf{1}^T\nu - \textbf{trace}(\Lambda(W + \textbf{diag}(\nu))) \\ &= \mathbf{1}^T\nu - \textbf{trace}(\Lambda W) - \textbf{trace}(\Lambda\textbf{diag}(\nu)) \\ &= -\textbf{trace}(\Lambda W) - \sum_{i=1}^{n} \underbrace{\nu_i(1 - \Lambda_{ii})}_{f(\nu_i, \Lambda_{ii}) :=} \end{aligned}$$

where $f(\nu_i, \Lambda_{ii})$ is linear in ν_i, hence unbounded below unless $1 - \Lambda_{ii} = 0$. Hence, the Lagrangian dual function $g(\Lambda)$ is:

$$g(\Lambda) = \inf_x L(X, \Lambda) = \begin{cases} -\textbf{trace}(\Lambda W) & \text{if } \Lambda_{ii} = 1, i = 1, ..., n \\ -\infty & \text{otherwise} \end{cases} \tag{7}$$

and the dual problem of (SDP1) is:

$$\begin{aligned} &\text{maximize} && -\textbf{trace}(WX) \\ &\text{subject to} && \Lambda \succeq 0, \\ &&& \Lambda_{ii} = 1,\ i = 1, ..., n \end{aligned}$$

Changing the sign of the objective function, the problem becomes (SDP2). Thus, (SDP2) is the dual problem of (SDP1). It is possible to find a primal-dual solution (ν^*, X^*) that satisfies the equality constraints and strictly satisfies the inequality constraints:

$$\begin{aligned} \nu^* &= \mathbf{1} \\ W + \textbf{diag}(\nu) = W + \mathbf{1} &\succ 0 \\ X^* = I_n &\succ 0 \\ X_{ii}^* &= 1,\ i = 1, ..., n \end{aligned}$$

Thus, the Slater condition is satisfied and the problems (SDP1) and (SDP2) have *strong duality*. Therefore, their optimal values provide the same lower bound to the original problem (P1).

References

[1] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.